

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-63300

(43) 公開日 平成10年(1998) 3月6日

(51) Int.Cl. ⁹	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 9/18			G 1 0 L 9/18	E
			9/14	G
H 0 3 M 7/30		9382-5K	H 0 3 M 7/30	B
H 0 4 B 14/04			H 0 4 B 14/04	Z

審査請求 未請求 請求項の数6 OL (全12頁)

(21) 出願番号 特願平8-220948

(22) 出願日 平成8年(1996) 8月22日

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 安永 和敏

神奈川県川崎市多摩区東三田3丁目10番1

号 松下技研株式会社内

(72) 発明者 森井 利幸

神奈川県川崎市多摩区東三田3丁目10番1

号 松下技研株式会社内

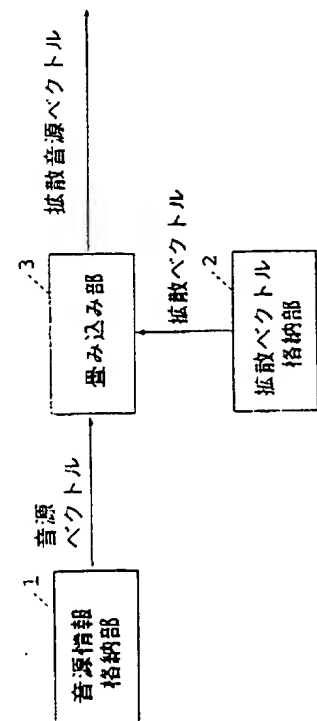
(74) 代理人 弁理士 滝本 智之 (外1名)

(54) 【発明の名称】 音声復号化装置及び音声符号化装置

(57) 【要約】

【課題】 音声情報の効率的な復号化及び符号化に使用される音声復号化装置及び音声符号化装置において、計算量が少なく、ROM容量が小さく、良好な合成音を提供することを目的とする。

【解決手段】 音声復号化装置では、少ないパルス本数により構成される音源ベクトルの生成情報を音源情報格納部1に格納し、ランダム数列、学習により求められた数列、知見により求められた数列のいずれか一つの数列を拡散ベクトル格納部2に格納するとともに、音源情報格納部1の音源情報と拡散ベクトル格納部2の拡散ベクトルとを畳み込み部3で畳み込むことにより、拡散音源ベクトルが作成されるように構成したものである。音声符号化装置では、拡散ベクトルとデジタルフィルタからのインパルス応答とを畳み込み部で畳み込むことにより新インパルス応答が作成されるように構成したものである。



【特許請求の範囲】

【請求項1】 音源情報を格納する音源情報格納部と、ランダム数列、学習により求められた数列、知見により求められた数列のいずれか一つの数列を格納する拡散ベクトル格納部と、前記音源情報を基に作成した音源ベクトルと前記拡散ベクトル格納部に格納された数列とを畳み込み、拡散音源ベクトルを出力する畳み込み部とを備えたことを特徴とする音声復号化装置。

【請求項2】 更に、過去の駆動音源を格納する適応符号帳と、前記適応符号帳から適応コードベクトルを読み出し、前記適応コードベクトルに前記適応符号帳の復号化ゲインを乗じて駆動音源適応符号帳成分を生成する機能、畳み込み部から拡散音源ベクトルを読み出し、前記拡散音源ベクトルに前記拡散音源ベクトルの復号化ゲインを乗じて駆動音源拡散音源ベクトル成分を生成する機能、前記駆動音源適応符号帳成分と前記駆動音源拡散音源ベクトル成分を加算して駆動音源を生成し、前記適応符号帳に出力する機能を持つ駆動音源生成部と、前記駆動音源をLPC合成する機能を持つLPC合成部とを備えたことを特徴とする請求項1記載の音声復号化装置。

【請求項3】 音源情報格納部に、代数的符号帳と同じコードベクトルを生成するための情報を格納することを特徴とする請求項1または2記載の音声復号化装置。

【請求項4】 ランダム数列、学習により求められた数列、知見により求められた数列のいずれか一つを格納する拡散ベクトル格納部と、デジタルフィルタと、前記拡散ベクトル格納部に格納された数列を前記デジタルフィルタのインパルス応答もしくは係数に畳み込み、前記デジタルフィルタの新インパルス応答もしくは新係数を出力する機能を有する畳み込み部とを備えたことを特徴とする音声符号化装置。

【請求項5】 更に、音源情報を格納する音源情報格納部を有し、畳み込み部が、前記音源情報を基に作成した音源ベクトルと拡散ベクトル格納部に格納された数列とを畳み込み、拡散音源ベクトルを出力する機能を併せ持つことを特徴とする請求項4記載の音声符号化装置。

【請求項6】 音源情報格納部に、代数的符号帳と同じコードベクトルを生成するための情報を格納することを特徴とする請求項5記載の音声符号化装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声復号化装置及び音声符号化装置における、音声情報の効率的な圧縮復号化及び符号化に関するものである。

【0002】

【従来の技術】デジタル携帯電話等の移動体通信の分野においては、加入者の増加に対処するため、低ビットレート（～8 kbps程度）の音声の圧縮符号化法が求められている。日本国内では、VSELPおよびPSI

ハーフレートのデジタル携帯電話の音声符号化標準方式として、それぞれ採用・実用化されている。国際的には、CS-ACELPという符号化方式が8 kbpsの国際標準音声符号化方式として採用され、ITU-T勧告G. 729となっている（1995年）。これらの音声符号化はいずれも、CELP方式（“CODE-EXCITED LINEAR PREDICTION (CELP): HIGH QUALITY SPEECH AT VERY LOW RATES”, Manfred R. Schroeder, Bishnu S. Atal, Proc., ICASSP'85, pp. 937-940に記載されている）を改良したものである。

【0003】ここで、CELP方式の基本的アルゴリズムについて説明する。CELP方式は、音声情報を音源情報と声道情報とに分離して符号化する方式で、音源情報については符号帳に格納された複数のコードベクトルのインデクスによって符号化し、声道情報についてはLPC（線形予測係数）を符号化するということと、音源情報符号化の際には声道情報を加味して入力音声と比較を行う方法（A-b-S: Analysis by Synthesis）を採用していることに特徴を有している。なおCELPでは一般に、入力音声のある時間間隔で区間（フレームと呼ばれる）ごとに分けてLPC分析を行い、フレームをさらに細かく分けた区間（サブフレームと呼ばれる）ごとに適応符号帳と確率的符号帳の音源探索が行われる。

【0004】図5は、従来のCELP型音声復号化装置の機能ブロック図である。パラメータ復号化部152は、従来のCELP型音声符号化装置（図6：後に説明する）から送られた音声符号（LPC符号、確率的符号帳のインデクス、確率的符号帳の符号化ゲイン、適応符号帳のインデクス、適応符号帳の符号化ゲイン）を、伝送部151を通して獲得する。次に、LPC符号を復号化して復号化LPC係数を得、確率的符号帳の符号化ゲインを復号化して確率的符号帳の復号化ゲインを得、適応符号帳の符号化ゲインを復号化して適応符号帳の復号化ゲインを得る。更に、確率的符号帳のインデクス、確率的符号帳の復号化ゲイン、適応符号帳のインデクス、適応符号帳の復号化ゲインを駆動音源生成部155へ出力し、復号化LPC係数をLPC合成部156へ出力する。

【0005】駆動音源生成部155は、まず、適応符号帳のインデクスに基づいた適応コードベクトルを適応符号帳153から読み出し、得られた適応コードベクトルに適応符号帳の復号化ゲインを乗じて駆動音源適応符号帳成分を得る。次に、確率的符号帳のインデクスに基づいた確率的コードベクトルを確率的符号帳154から読み出し、得られた確率的コードベクトルに確率的符号帳の復号化ゲインを乗じて駆動音源確率的符号帳成分を得る。更に、駆動音源適応符号帳成分と駆動音源確率的符号帳成分を加算して駆動音源を得、得られた駆動音源をLPC合成部156と適応符号帳155へ出力する。

ここで、適応符号帳153内の古いコードベクトルは、駆動音源生成部155から入力された上記駆動音源で更新される。

【0006】LPC合成部156は、駆動音源生成部155で得られた駆動音源に対し、パラメータ復号化部152より得た復号化LPC係数をもとにLPC合成を行い、その出力をデジタルの出力音声157として出力部位へ送る。

【0007】図6は、従来のCELP型音声符号化装置の機能ブロック図である。LPC分析部112は、まず、デジタルの入力音声111内のあるフレームに対して自己相関分析と線形予測分析を行うことによってLPC係数を算出し、そのLPC係数を量子化してLPC符号を得てパラメータ符号化部123へ出力し、LPC符号を復号化して復号化LPC係数を得、次に、ピッチ強調や高域強調などの特性を持つ聴感重み付けフィルタのインパルス応答を求めて聴感重み付け部113へ出力するとともに、聴感重み付けLPC合成フィルタのインパルス応答を求めて、聴感重み付けLPC逆順合成部A114、聴感重み付けLPC合成部A116、聴感重み付けLPC逆順合成部B119、聴感重み付けLPC合成部B121へ出力する。

【0008】聴感重み付け部113は、入力された音声データに対し、サブフレーム毎に聴感重み付けフィルタリングを行い、その出力結果から聴感重み付けLPC合成フィルタのゼロ入力応答を差し引いて、適応符号帳の音源探索時に参照するターゲット信号を求め、聴感重み付けLPC逆順合成部A114および減算部118へ出力する。

【0009】聴感重み付けLPC逆順合成部A114は、聴感重み付け部113で得られたターゲット信号を時間逆順化し、得られた逆順化信号をLPC分析部112より与えられたインパルス応答を係数に持つ聴感重み付けLPC合成フィルタで合成し、その出力信号を再度時間逆順化して、ターゲット信号の時間逆合成出力として比較部A117へ出力する。

【0010】適応符号帳115は、適応符号帳更新部124により受けた過去の駆動音源を格納しており、その過去の駆動音源情報は、聴感重み付けLPC合成部A116、比較部A117、適応符号帳更新部124により、適応コードベクトルとして参照される。

【0011】聴感重み付けLPC合成部A116は、適応符号帳115から適応コードベクトルを読み出し、読み出した適応コードベクトルに対し、LPC分析部112より得たインパルス応答を係数に持つ聴感重み付けLPC合成フィルタで合成し、その結果を比較部A117へ出力する。

【0012】比較部A117は、まず、適応符号帳115から直接読み出した適応コードベクトルと、聴感重み

付けLPC逆順合成部A114で求めたターゲット信号の時間逆合成出力との内積の2乗値を求め、次に、聴感重み付けLPC合成部A116から受けた、適応コードベクトルに聴感重み付けLPC合成を施した信号のパワを求め、そして、上記内積の2乗値をこのパワで割算することによって適応符号帳探索の基準値を求め、その基準値が最も大きくなるときに読み出した適応コードベクトルのインデクスと、そのコードベクトルに乗じる最適ゲインを算出して、減算部118およびパラメータ符号化部123へ出力する。この一連の処理を適応符号帳の探索という。

【0013】減算部118は、聴感重み付け部113で得たターゲット信号から、適応符号帳探索によって探索されたコードベクトルを聴感重み付けLPC合成した出力信号にゲインを乗じて得られた信号を減算し、その減算結果を確率的符号帳の探索の際に参照するターゲット信号として聴感重み付けLPC逆順合成部B119へ出力する。

【0014】聴感重み付けLPC逆順合成部B119は、減算部118において生成された確率的符号帳の音源探索用のターゲット信号を時間逆順化し、それを聴感重み付けLPC合成し、その出力信号を再度時間逆順化して、確率的符号帳の音源探索用ターゲット信号の時間逆合成出力を得て、比較部B122へ出力する。

【0015】確率的符号帳120は、複数のコードベクトルを格納しており、これらのコードベクトルは、聴感重み付けLPC合成部B121、比較部B122、適応符号帳更新部124により、確率的コードベクトルとして参照される。

【0016】聴感重み付けLPC合成部B121は、確率的符号帳120から読み出した確率的コードベクトルに対し、LPC分析部112から得たインパルス応答を係数に持つ聴感重み付けLPC合成フィルタで合成し、その合成信号を比較部B122へ出力する。

【0017】比較部B122は、(数1)に示すように、まず、確率的符号帳120より直接読み出したi番目の確率的コードベクトル $V(i, n)$ と、聴感重み付けLPC逆順合成部B119で求めたターゲット信号の時間逆合成出力 $r(n)$ との内積の2乗値を求め、次に、聴感重み付けLPC合成部B121より受けた合成信号 $S(i, n)$ のパワを計算し、そして上記内積の2乗値をこのパワで割算することによって確率的符号帳探索の基準値 $std(i)$ を求め、その基準値が最も大きくなるときに読み出した確率的コードベクトルの番号を表すインデクスと、その確率的コードベクトルに乗じる最適ゲインを算出して、パラメータ符号化部123へ出力する。この一連の処理を確率的符号帳の探索という。

【0018】

【数1】

$$\text{std}(i) = \frac{\left| \sum_{n=0}^{N-1} V(i, n) \cdot r(n) \right|^2}{\sum_{n=0}^{N-1} S^2(i, n)}$$

$\left\{ \begin{array}{ll} \text{std}(i) & : \text{確率的符号帳探索の基準値} \\ i & : \text{コードベクトルのインデックス} \\ N & : \text{サブフレーム長} \\ n & : \text{ベクトルの成分番号} (0 \leq n < N) \\ V(i, n) & : \text{i番目の確率的コードベクトル} \\ S(i, n) & : \text{V(i, n)を聴感重み付けしLPC合成したベクトル} \\ r(n) & : \text{確率的符号帳探索用ターゲット信号の時間逆合成出力ベクトル} \end{array} \right.$

【0019】パラメータ符号化部123は、まず、比較部A117から得られた適応コードベクトルに乗じる最適ゲインと、比較部B122で得られた確率的コードベクトルに乗じる最適ゲインの符号化を行って、適応符号帳の符号化ゲインと確率的符号帳の符号化ゲインをそれぞれ得、次に、得られた適応符号帳の符号化ゲインと確率的符号帳の符号化ゲイン復号化して、適応符号帳の復号化ゲインと確率的符号帳の復号化ゲインをそれぞれ得、さらに、適応符号帳の符号化ゲイン、確率的符号帳の符号化ゲイン、適応符号帳のインデックス、確率的符号帳のインデックス、LPC符号を伝送部125へ出力し、適応符号帳の復号化ゲイン、確率的符号帳の復号化ゲイン、適応符号帳のインデックス、確率的符号帳のインデックスを適応符号帳更新部124へ出力する。

【0020】適応符号帳更新部124は、パラメータ符号化部123からの入力を受けて、まず、適応符号帳のインデックスに基づいた適応コードベクトルを適応符号帳115から読みだし、読み出した適応コードベクトルに適応符号帳の復号化ゲインを乗じて駆動音源適応符号帳成分を得、次に、確率的符号帳のインデックスに基づいた確率的コードベクトルを確率的符号帳120から読みだし、読み出した確率的コードベクトルに確率的符号帳の復号化ゲインを乗じて駆動音源確率的符号帳成分を得、駆動音源適応符号帳成分と駆動音源確率的符号帳成分を加算して駆動音源を得、得られた駆動音源を適応符号帳115へ出力する。ここで、適応符号帳115内の古いコードベクトルは、適応符号帳更新部124から入力された上記駆動音源で更新される。

【0021】CELPで用いる上記確率的符号帳としては、雑音符号帳や代数的符号帳（“8KBIT/S ACELP CODING OF SPEECH WITH 10 MS SPEECH-FRAME: A CANDIDATE FOR CCITT STANDARDIZATION”：R. Salami, C. Laflamme, J.-P. Adoul, ICASSP'94, pp. II-97~II-100, 1994に記載されている）などがある。それぞれを簡単に説明する。

【0022】「雑音符号帳」は最も古典的な符号帳で、乱数から作成したランダム数列を格納したものである。コードブックの性質がランダムなので、質の高い合成音を得ることができる。しかし、全コードベクトルをあらかじめ格納しておかなければならないので、大きなROM容量が必要となる。また、符号帳探索は、全てのコードベクトルに対して聴感重み付けしLPC合成を行うこと

になるので、多くの計算量を必要としてしまう。

【0023】「代数的符号帳」の個々のコードベクトルは、大きさが1（振幅は、+1か-1）の4本のパルスで構成され、それぞれのパルスの位置はインデックスの演算により決定されることを特徴としている。したがって、符号帳用のROMを必要としない。しかし、少数のパルスで構成されることから、合成音（特に無声部）の音質が悪くなってしまう。そして代数的符号帳の最大の特徴は、少量の計算量で符号帳探索ができるということである。代数的符号帳における上記最大の特徴は、V(i, n)中には大きさ1のパルスが4本しか存在していないので、(数1)の分子を、r(n)中の4サンプルの値を加算し（加算：3回）、その加算結果を2乗（乗算：1回）することで計算できることと、前もって求めておいた聴感重み付けLPC合成フィルタのインパルス応答の自己相関行列をRAMに格納しておく、(数1)の分母を、たかだか15回の加算（4重ループの性質や自己相関行列の対称性を利用すると、実際の計算量はさらに少なくなる。）によって計算することができることによるものである。

【0024】

【発明が解決しようとする課題】雑音符号帳を用いると、ランダムな性質を持つ音源で聴感重み付けLPC合成フィルタを駆動できるので合成音の音質は高くなるが、(数1)の分子および分母の計算量が多くなるので符号帳探索の計算量が多くなってしまい、また、全コードベクトルをあらかじめ格納しておくためにROM容量が大きくなってしまう。

【0025】一方、代数的符号帳を用いると、(数1)の分子および分母の計算量が少なくなるので符号帳探索の計算量が少なくなり、コードベクトルをそのまま格納する必要がないのでROM容量が小さくなるが、少数本のパルスによって聴感重み付けLPC合成フィルタを駆動することになるため合成音（特に無声部）の音質が悪くなってしまう。

【0026】本発明は、符号帳探索のための計算量が少なく、ROM容量が小さく、良好な合成音を提供することが可能であるという有利な作用を併せ持つ音声復号化装置及び音声符号化装置を構成することを目的とする。

【0027】

【課題を解決するための手段】一の課題を解決するため

に本発明の音声復号化装置は、音源情報を格納する音源情報格納部と、ランダム数列、学習により求められた数列、知見により求められた数列のいずれか一つの数列を格納する拡散ベクトル格納部と、前記音源情報を基に作成した音源ベクトルと前記拡散ベクトル格納部に格納された数列とを畳み込み、拡散音源ベクトルを出力する畳み込み部とを備え、少ないパルス本数により構成される音源ベクトルの生成情報を音源情報格納部に格納し、前記いずれか一つの数列を拡散ベクトル格納部に格納するとともに、拡散音源ベクトルが、音源情報格納部内の音源情報と拡散ベクトル格納部内の拡散ベクトルとを畳み込むことにより作成されるように構成したものである。

【0028】これにより、合成音（特に無声部）の音質に優れ、雑音符号帳を備える必要のない、ROM容量の小さな音声復号化装置が得られる。

【0029】また、本発明の音声符号化装置は、ランダム数列、学習により求められた数列、知見により求められた数列のいずれか一つを格納する拡散ベクトル格納部と、デジタルフィルタと、前記拡散ベクトル格納部に格納された数列を前記デジタルフィルタのインパルス応答もしくは係数に畳み込み、前記デジタルフィルタの新インパルス応答もしくは新係数を出力する機能を有する畳み込み部とを備え、新たなフィルタ特性を決定づける新インパルス応答もしくは新係数が、デジタルフィルタ内のインパルス応答もしくは係数と拡散ベクトル格納部内の拡散ベクトルとを畳み込むことにより作成されるように構成したものである。

【0030】これにより、少ないパルス本数により構成される音源ベクトルの探索と同じように、少ない計算量で符号帳探索が可能な音声符号化装置が得られる。

【0031】

【発明の実施の形態】本発明の請求項1に記載の発明は、音源情報を格納する音源情報格納部と、ランダム数列、学習により求められた数列、知見により求められた数列のいずれか一つの数列を格納する拡散ベクトル格納部と、前記音源情報を基に作成した音源ベクトルと前記拡散ベクトル格納部に格納された数列とを畳み込み、拡散音源ベクトルを出力する畳み込み部とを備えたことを特徴とする音声復号化装置で、音声情報の伝送に関する課題解決に関わるものであり、雑音符号帳や代数的符号帳等の確率的符号帳を備える必要がなく、ROM容量が小さくなり、音源情報のパルス本数が少なく、計算量が少なくなるという作用を有する。

【0032】本発明の請求項2に記載の発明は、過去の駆動音源を格納する適応符号帳と、前記適応符号帳から適応コードベクトルを読み出し、前記適応コードベクトルに前記適応符号帳の復号化ゲインを乗じて駆動音源適応符号帳成分を生成する機能、畳み込み部から拡散音源ベクトルを読み出し、前記拡散音源ベクトルに前記拡散音源ベクトルの復号化ゲインを乗じて駆動音源拡散音源

ベクトル成分を生成する機能、前記駆動音源適応符号帳成分と前記駆動音源拡散音源ベクトル成分を加算して駆動音源を生成し、前記適応符号帳に出力する機能を持つ駆動音源生成部と、前記駆動音源をLPC合成する機能を持つLPC合成部とを備えたことを特徴とする請求項1記載の音声復号化装置であり、音声情報のうち声道情報のLPCを符号化でき、良好な合成音が生成され、雑音符号帳や代数的符号帳等の確率的符号帳を備える必要がなく、ROM容量が小さくなり、更に音源情報のパルス本数が少なく、計算量が少なくなるという作用を有する。

【0033】本発明の請求項3に記載の発明は、音源情報格納部に、代数的符号帳と同じコードベクトルを生成するための情報を格納することを特徴とする請求項1または2記載の音声復号化装置であり、音源情報のパルス本数が少なく、計算量が少なくなるという作用を有する。

【0034】本発明の請求項4に記載の発明は、ランダム数列、学習により求められた数列、知見により求められた数列のいずれか一つを格納する拡散ベクトル格納部と、デジタルフィルタと、前記拡散ベクトル格納部に格納された数列を前記デジタルフィルタのインパルス応答もしくは係数に畳み込み、前記デジタルフィルタの新インパルス応答もしくは新係数を出力する機能を有する畳み込み部とを備えたことを特徴とする音声符号化装置で、音声情報の伝送に関する課題解決に関わるものであり、新たなフィルタ特性を決定づける新インパルス応答もしくは新係数が、デジタルフィルタ内のインパルス応答もしくは係数と拡散ベクトル格納部内の拡散ベクトルとを畳み込むことにより作成され、復号化装置における復号化との対応が適切となるような符号化を行うために必要なデジタルフィルタが形成されるという作用を有する。

【0035】本発明の請求項5に記載の発明は、音源情報を格納する音源情報格納部を有し、畳み込み部が、前記音源情報を基に作成した音源ベクトルと拡散ベクトル格納部に格納された数列とを畳み込み、拡散音源ベクトルを出力する機能を併せ持つことを特徴とする請求項4記載の音声符号化装置であり、新たなフィルタ特性を決定づける新インパルス応答もしくは新係数が、デジタルフィルタ内のインパルス応答もしくは係数と拡散ベクトル格納部内の拡散ベクトルとを畳み込むことにより作成され、復号化装置における復号化との対応が最適となるような符号化を行うために必要なデジタルフィルタが形成されるとともに、雑音符号帳や代数的符号帳等の確率的符号帳を備える必要がなく、ROM容量が小さくなり、更に音源情報のパルス本数が少なく、計算量が少なくなるという作用を有する。

【0036】本発明の請求項6に記載の発明は、音源情報格納部に、代数的符号帳と同じコードベクトルを生成

するための情報を格納することを特徴とする請求項5記載の音声復号化装置であり、音源情報のパルス本数が少なく、計算量が少なくなるという作用を有する。

【0037】以下、本発明の実施の形態について、図1から図4を用いて説明する。

(実施の形態1) 本発明の実施の形態を、図1と図2を用いて説明する。

【0038】図1は本発明による音声復号化装置の主要部のブロック図を示す。図1において、1は音源情報格納部、2は拡散ベクトル格納部、3は畳み込み部である。音源情報格納部1は、例えば代数的符号帳と同じコードベクトルの生成情報等の音源情報を格納し、その音源情報を音源ベクトルとして出力する。拡散ベクトル格納部2は、例えばランダム数列等の拡散情報を格納し、その拡散情報を拡散ベクトルとして出力する。畳み込み部3は音源ベクトルと拡散ベクトルを入力して畳み込みを行い、拡散音源ベクトルとして出力する。

【0039】図2は、図1の音声復号化装置の主要部を従来の確率的符号帳部に代えて用いたCELP型の音声復号化装置を示す。音源情報格納部1は代数的符号帳と同じコードベクトルの生成情報を格納し、拡散ベクトル格納部2はランダム数列を拡散ベクトルとして格納している。パラメータ復号化部4は、まず、符号化装置が生成した音声符号(LPC符号、拡散音源インデクス、拡散音源ベクトルの符号化ゲイン、適応符号帳のインデクス、適応符号帳の符号化ゲイン)を伝送部5を通して獲得し、LPC符号を復号化して復号化LPC係数を得る。次に拡散音源ベクトルの符号化ゲインを復号化して拡散音源ベクトルの復号化ゲインを得、適応符号帳の符号化ゲインを復号化して適応符号帳の復号化ゲインを得る。更に、拡散音源インデクスを畳み込み部3へ出力し、適応符号帳のインデクス及び復号化ゲイン、拡散音源ベクトルの復号化ゲインを駆動音源生成部7へ出力し、復号化LPC係数をLPC合成部8へ出力する。

【0040】畳み込み部3は、まず、拡散音源インデクスに基づいた音源ベクトルを音源情報格納部1から読み出し、次に、拡散ベクトル格納部2に格納されている拡散ベクトルを読み出し、更に、読み出した音源ベクトルと拡散ベクトルとの畳み込みを行って拡散音源ベクトルを生成し駆動音源生成部7へ出力する。駆動音源生成部7は、まず、適応符号帳のインデクスに基づいた適応コードベクトルを適応符号帳6から読み出し、得られた適応コードベクトルに適応符号帳の復号化ゲインを乗じて駆動音源適応符号帳成分を得、畳み込み部3より得られた拡散音源ベクトルに拡散音源ベクトルの復号化ゲインを乗じて駆動音源拡散音源ベクトル成分を得る。次に、得られた駆動音源適応符号帳成分と駆動音源拡散音源ベクトル成分を加算して駆動音源を得、得られた駆動音源をLPC合成部8と適応符号帳6へ出力する。ここで、適応符号帳6内の適応コードベクトルは、駆動音源生成

部7から入力された上記駆動音源で更新される。LPC合成部8は、駆動音源生成部7から入力される駆動音源に対し、パラメータ復号化部4から得た復号化LPC係数を持つLPC合成フィルタで合成し、出力音声9を得る。

【0041】なお、本実施の形態において拡散ベクトル格納部2は、ランダム数列を拡散ベクトルとして格納しているので、上記駆動音源拡散音源ベクトル成分がランダムな性質をもつことになり、出力音声(特に無声部)の音質を向上することができる。また、拡散音源ベクトルは音源情報格納部1の音源情報と拡散ベクトル格納部2の拡散ベクトルにより作成できるので、ROM容量は小さくなっている。

【0042】なお、本実施の形態では、音源情報格納部1に代数的符号帳と同じコードベクトルの生成情報を格納した例で説明したが、その他の符号帳の生成情報もしくはその他の符号帳自身を格納した場合についても同様に実施可能である。また、本実施の形態では、拡散ベクトル格納部2にランダム数列を格納した例で説明したが、その他の学習により求められた数列、もしくは知見により求められた数列を用いる場合についても同様に実施可能である。

【0043】なお、本実施の形態における音声復号化装置はCELP型としたが、VOCODER型等、その他の音声復号化装置においても適用が可能である。

【0044】(実施の形態2) 本発明の実施の形態を、図3と図4を用いて説明する。

【0045】図3は本発明による音声符号化装置の主要部のブロック図を示す。図3において、11は拡散ベクトル格納部、12はディジタルフィルタ、13は畳み込み部、14は音源情報格納部である。拡散ベクトル格納部11は、例えばランダム数列等の拡散情報を格納し、その拡散情報を拡散ベクトルとして出力する。ディジタルフィルタ12は、入力信号をフィルタリングして出力するとともに、フィルタ自身の特性を決めるフィルタ情報を出力するもので、図3においてはフィルタ情報としてインパルス応答もしくは係数の出力を示す。音源情報格納部14は、例えば代数的符号帳と同じコードベクトルの生成情報等の音源情報を格納し、その音源情報を音源ベクトルとして出力する。畳み込み部13は、図3

(a)では、拡散ベクトルとインパルス応答もしくは係数を入力して畳み込みを行い、新インパルス応答もしくは新係数を出力し、図3(b)では、図3(a)の機能に加えて、音源ベクトルと拡散ベクトルを入力して畳み込みを行い、拡散音源ベクトルとして出力する。

【0046】図4は、図3の音声符号化装置の主要部を従来の確率的符号帳部に代えて用いたCELP型の音声符号化装置を示す。音源情報格納部14は代数的符号帳と同じコードベクトルの生成情報を格納し、拡散ベクトル格納部11はランダム数列を拡散ベクトルとして格納

している。15は入力音声で、デジタルの入力音声データである。LPC分析部16は、入力音声15におけるあるフレームに対して自己相関分析と線形予測分析を行うことによってLPC係数を算出し、そのLPC係数を量子化してLPC符号を得てパラメータ符号化部17へ出力し、LPC符号を復号化して復号化LPC係数を得る。次に、ピッチ強調や高域強調などの特性を持つ聴感重み付けフィルタのインパルス応答を求めて聴感重み付け部18へ出力するとともに、聴感重み付けLPC合成フィルタのインパルス応答を求めて、聴感重み付けLPC逆順合成部A19および聴感重み付けLPC合成部A20へ出力する。聴感重み付けLPC逆順合成部A19と聴感重み付けLPC合成部A20はいずれもデジタルフィルタを含む。

【0047】聴感重み付け部18は、入力音声データに対し、サブフレーム毎に聴感重み付けフィルタリングを行い、その出力結果から聴感重み付けLPC合成フィルタのゼロ入力応答を差し引いて適応符号帳の音源探索時に参照するターゲット信号を求め、聴感重み付けLPC逆順合成部A19および減算部18へ出力する。聴感重み付けLPC逆順合成部A19は、聴感重み付け部18から入力されるターゲット信号を時間逆順化し、その逆順化信号をLPC分析部16より与えられたインパルス応答を係数に持つ聴感重み付けLPC合成フィルタで合成し、その出力信号を再度時間逆順化してターゲット信号の時間逆合成出力を得て比較部A21へ出力する。

【0048】適応符号帳22は、適応符号帳更新部23により受けた過去の駆動音源を格納しており、その過去の駆動音源情報は、聴感重み付けLPC合成部A20、比較部A21、適応符号帳更新部23により、適応コードベクトルとして参照される。聴感重み付けLPC合成部A20は、まず、適応符号帳22から適応コードベクトルを読み出し、読み出した適応コードベクトルに対し、LPC分析部16より得たインパルス応答を係数に持つ聴感重み付けLPC合成フィルタで合成し、その結果を比較部A21へ出力する。次に、聴感重み付けLPC合成フィルタのインパルス応答を畳み込み部13へ出力する。

【0049】比較部A21は、まず、適応符号帳22から直接読み出した適応コードベクトルと、聴感重み付けLPC逆順合成部A19で求めたターゲット信号の時間逆合成出力との内積の2乗値を求める。次に、聴感重み付けLPC合成部A20から受けた適応コードベクトルに聴感重み付けLPC合成を施した信号のパワを求め、上記内積の2乗値をこのパワで割算することによって適応コードベクトル探索の基準値を求める。その基準値が最も大きくなるときに読み出した適応コードベクトルのインデクスと、その適応コードベクトルに乗じる最適ゲインを算出し、減算部24およびパラメータ符号化部17へ出力する。

【0050】減算部24は、聴感重み付け部18で得られたターゲット信号から、適応符号帳探索において探索された適応コードベクトルを聴感重み付けLPC合成し、合成信号に比較部A21で求めた適応符号帳の最適ゲインを乗じて得られた信号を減算する。その減算結果を、拡散音源ベクトルの探索の際に参照するターゲット信号として聴感重み付けLPC逆順合成部B25へ出力する。

【0051】拡散ベクトル格納部11は、ランダム数列を格納している。畳み込み部13は、聴感重み付けLPC合成部A20から受けた聴感重み付けLPC合成フィルタのインパルス応答と拡散ベクトル格納部11から読み出した拡散ベクトルとの畳み込み演算を行い、その演算結果を新インパルス応答として、聴感重み付けLPC逆順合成部B25と聴感重み付けLPC合成部B26へ出力する。聴感重み付けLPC逆順合成部B25と聴感重み付けLPC合成部B26はいずれもデジタルフィルタを含む。

【0052】聴感重み付けLPC逆順合成部B25は、減算部24で生成された拡散音源ベクトル探索時のターゲット信号を時間逆順化し、その逆順化した信号を、畳み込み部13から得た新インパルス応答を係数に持つ聴感重み付けLPC合成フィルタで合成する。その出力信号を再度時間逆順化し、ターゲット信号の時間逆合成出力として比較部B27へ出力する。

【0053】音源情報格納部14は、代数的符号帳と同じコードベクトルの生成情報を格納しているので、少数（ここでは4本）のパルスよりなるコードベクトルを生成することができる。生成したコードベクトルは、聴感重み付けLPC合成部B26、比較部B27、畳み込み部13により、音源ベクトルとして参照される。聴感重み付けLPC合成部B26は、音源情報格納部14より読み出した音源ベクトルに対し、畳み込み部13から得た新インパルス応答を係数に持つ聴感重み付けLPC合成フィルタで合成し、その合成信号の自己相関行列を求めて比較部B27へ出力する。

【0054】比較部B27は、まず、音源情報格納部14より直接読み出した音源ベクトルと、聴感重み付けLPC逆順合成部B25で求めたターゲット信号の時間逆合成出力との内積の2乗値を求める。次に、聴感重み付けLPC合成部B26より受けた自己相関行列を参照して、音源ベクトルに聴感重み付けLPC合成を施した信号のパワを求め、そして上記内積の2乗値をこのパワで割算することによって拡散音源ベクトル探索の基準値を求める。その基準値が最も大きくなるときに読み出した音源ベクトルの番号を表す拡散音源インデクスと、その拡散音源ベクトルに乗じる最適ゲインを算出して、パラメータ符号化部17へ出力する。

【0055】パラメータ符号化部17は、まず、比較部A21で得られた適応コードベクトルに乗じる最適ゲイ

ンと、比較部B27で得られた拡散音源ベクトルに乗じる最適ゲインの符号化を行って、適応符号帳の符号化ゲインと拡散音源ベクトルの符号化ゲインをそれぞれ得る。次に、得られた適応符号帳の符号化ゲインと拡散音源ベクトルの符号化ゲインを復号化して、適応符号帳の復号化ゲインと拡散音源ベクトルの復号化ゲインをそれぞれ得る。更に、適応符号帳の符号化ゲイン、拡散音源ベクトルの符号化ゲイン、LPC符号、適応符号帳のインデックス、拡散音源インデックスを伝送部28へ出力し、適応符号帳の復号化ゲインと、適応符号帳のインデックスとを適応符号帳更新部23へ出力し、拡散音源ベクトルの復号化ゲイン、拡散音源インデックスを畳み込み部13へ出力する。

【0056】畳み込み部13は、まず、拡散音源インデックスに基づいた音源ベクトルを音源情報格納部14から読み出し、次に、拡散ベクトル格納部11に格納されている拡散ベクトルを読み出す。次に、読み出した音源ベクトルと拡散ベクトルとの畳み込み演算を行って拡散音源ベクトルを生成し、拡散音源ベクトルに復号化ゲインを乗じて駆動音源拡散音源ベクトル成分を得て適応符号帳更新部23へ出力する。適応符号帳更新部23は、まず、パラメータ符号化部17から受けた適応符号帳のインデックスに基づいた適応コードベクトルを適応符号帳22から読み出し、得られた適応コードベクトルに適応符号帳の復号化ゲインを乗じて駆動音源適応符号帳成分を得る。次に、得られた駆動音源適応符号帳成分と畳み込み部13から入力された駆動音源拡散音源ベクトル成分を加算して駆動音源を生成し、生成した駆動音源を適応符号帳22へ出力する。ここで、適応符号帳22内の古いコードベクトルは、適応符号帳更新部23から入力された上記駆動音源で更新される。

【0057】なお、本実施の形態では、音源情報格納部14に代数的符号帳と同じコードベクトルの生成情報を格納した例で説明したが、その他の符号帳の生成情報もしくはその他の符号帳自身を格納した場合についても同様に実施可能である。また、本実施の形態では、拡散ベクトル格納部11にランダム数列を格納した例で説明したが、その他の学習により求められた数列、もしくは知見により求められた数列を用いる場合についても同様に実施可能である。

【0058】なお、本実施の形態における音声符号化装置はCELP型としたが、VOCODER型等、その他の音声符号化装置においても適用が可能である。

【0059】

【発明の効果】以上のように本発明によれば、符号帳探索のための計算量が少なくて済むとともに、ROM容量

が小さくて済み、更に良好な合成音を提供できるという有利な効果が得られる。

【図面の簡単な説明】

【図1】本発明の一実施の形態による音声復号化装置の主要部を示すブロック図

【図2】本発明の一実施の形態によるCELP型の音声復号化装置を示すブロック図

【図3】(a)本発明の一実施の形態による音声符号化装置の主要部を示すブロック図

(b)本発明の一実施の形態による音声符号化装置の主要部を示すブロック図

【図4】本発明の一実施の形態によるCELP型の音声符号化装置を示すブロック図

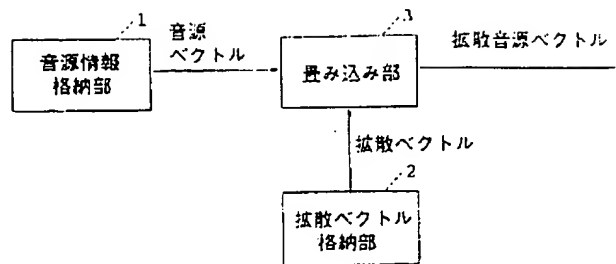
【図5】従来のCELP型音声復号化装置を示すブロック図

【図6】従来のCELP型音声符号化装置を示すブロック図

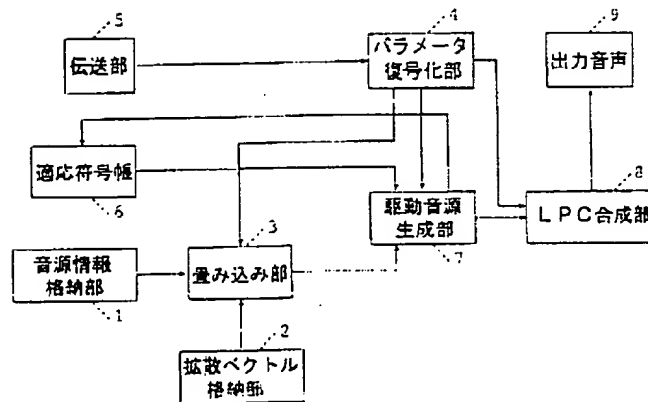
【符号の説明】

- 1 音源情報格納部
- 2 拡散ベクトル格納部
- 3 畳み込み部
- 4 パラメータ復号化部
- 5 伝送部
- 6 適応符号帳
- 7 駆動音源生成部
- 8 LPC合成部
- 9 出力音声
- 11 拡散ベクトル格納部
- 12 デジタルフィルタ
- 13 畳み込み部
- 14 音源情報格納部
- 15 入力音声
- 16 LPC分析部
- 17 パラメータ符号化部
- 18 聴感重み付け部
- 19 聴感重み付けLPC逆順合成部A
- 20 聴感重み付けLPC合成部A
- 21 比較部A
- 22 適応符号帳
- 23 適応符号帳更新部
- 24 減算部
- 25 聴感重み付けLPC逆順合成部B
- 26 聴感重み付けLPC合成部B
- 27 比較部B
- 28 伝送部

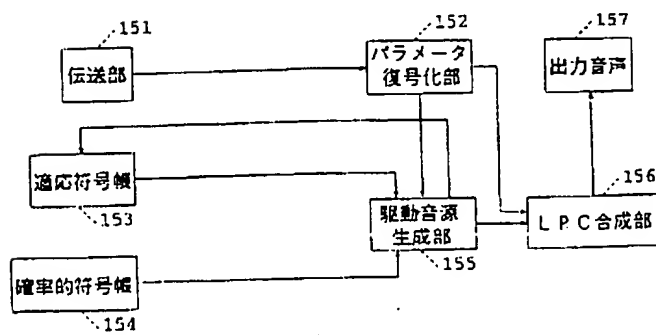
【図1】



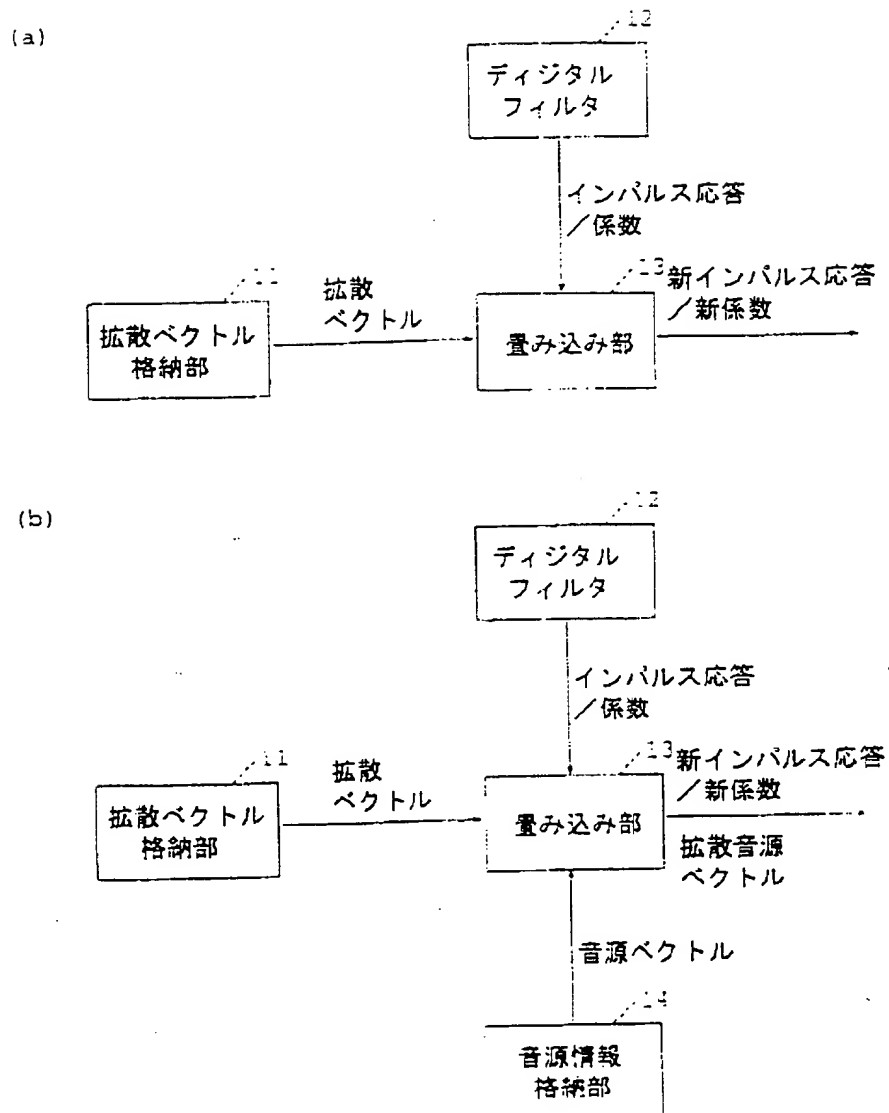
【図2】



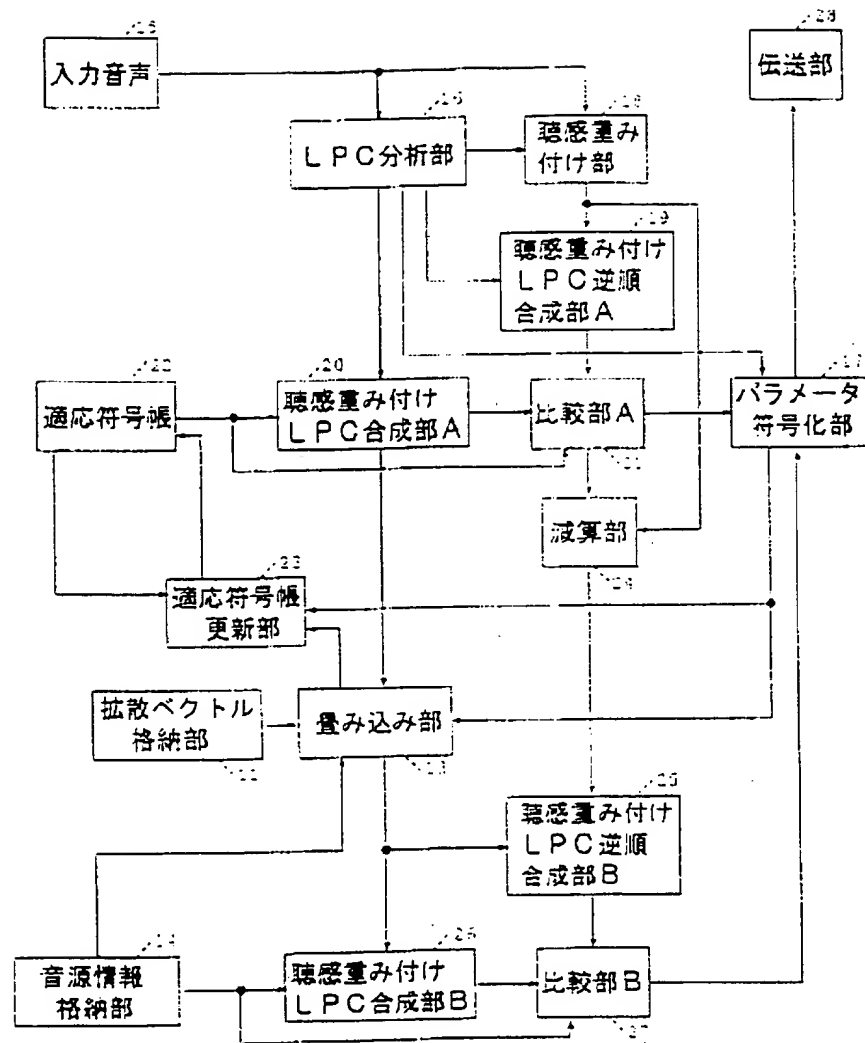
【図5】



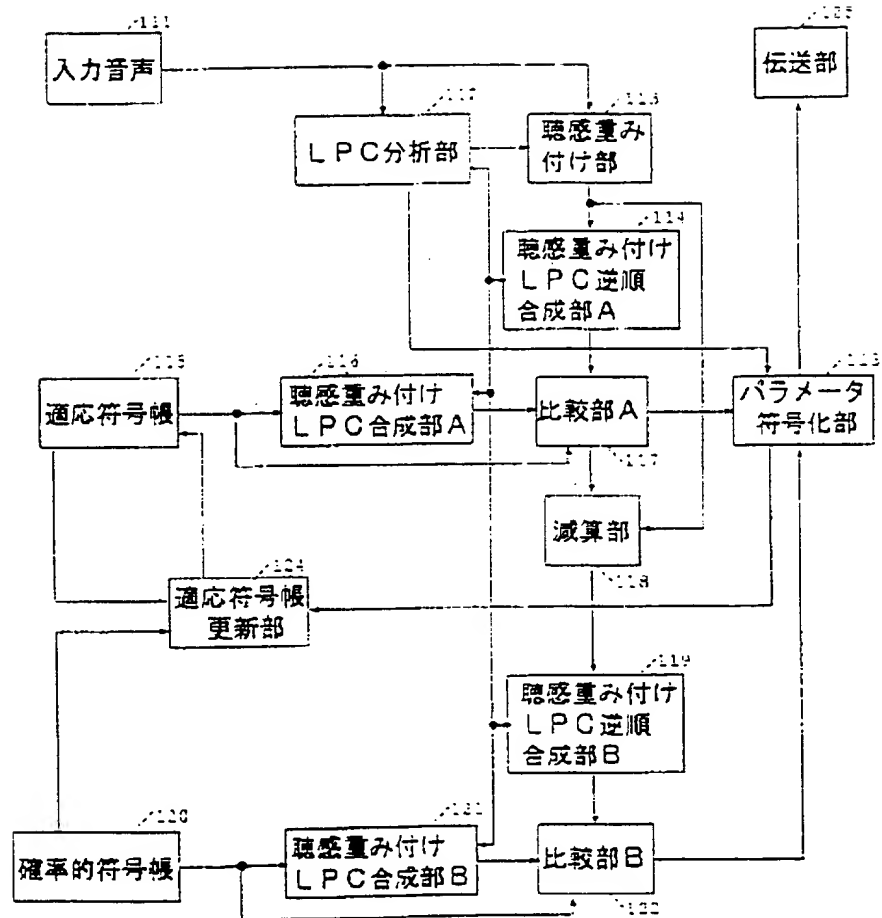
【図3】



【圖 4】



【図6】



(19) Japanese Patent Office (JP)

(12) Laid-open Patent Publication (A)

(11) Patent Application Laid-open No. Hei-10-
63300

5 (43) Laid open on March 6, 1998

(51) International Classification

G10L 9/18 9/14. H03M 7/30 H04B 14/04

Technical Indication points

Examination request, not requested, Number of
10 claims : 6 (OL) (12 pages in all)

(21) File number : Patent Application No. Hei-8-
220948

(22) Date of filing: August 22, 1996

(71) Applicant

15 000005821

Matsushita Electric Industries Co., Ltd.

1006, Kadoma, Ohaza, Kadomashi, Osaka-fu

(72) Inventor

Kazutoshi Yasunaga

20 c/o Matsushita Giken Co., Ltd. 10-1, 3-
chome, Higashimita, Tama-ku, Kawasakishi,
Kanagawa-Ken,

(73) Inventor

Toshiyuki Morii

25 c/o Matsushita Giken Co., Ltd. 10-1, 3-
chome, Higashimita, Tama-ku, Kawasakishi,
Kanagawa-Ken,

(74) Patent attorney

Tomoyuki Takimoto (et al)

(54) [Title of invention] Speech decoder
and speech coder

5 (57) [Abstract]

[Object] It is the object to provide
good quality synthetic sounds in a speech
decoder and a speech coder, having minimal
calculations and a small ROM capacity,
10 which are used to efficiently decode and
code speech information..

[Overcoming Means] In a speech decoder,
generated information of excitation
vectors, which are constituted by a limited
15 number of pulses, is stored in an
excitation information storing section 1,
either matrix of random matrices, matrices
obtained by learning and matrices obtained
by knowledge are stored in a diffused
20 vector storing section 2, and at the same
time, excitation information in the
excitation information storing section 1
and diffusion vectors in the diffusion
vector storing section 2 are convoluted by
25 a convolution section 3, whereby diffusion
excitation vectors are generated. The
speech coder is constituted so that

diffusion vectors and impulse responses from a digital filter are convoluted by the convolution section, whereby new impulse responses are generated.

5 [SCOPE OF CLAIMS FOR PATENT]

[Claim 1] A speech decoder comprising an excitation information storing section for storing excitation information; a diffusion vector storing section for storing either
10 matrix of random matrices, matrices obtained by learning, and matrices obtained by knowledge; and a convolution section, for outputting diffusion excitation vectors, which convolutes excitation vectors generated on the basis of
15 said excitation information and matrices stored in said diffusion vector storing section.

[Claim 2] A speech decoder, as set forth in Claim 1, further comprising an adaptive code book for storing the previously
20 stored excitation vectors; an excitation vector generator having a function to generate excitation vector adaptive code book components by reading adaptive code vectors from said adaptive code book and multiplying said
25 adaptive code vectors by decoder gains of said adaptive code book, a function to generate excitation vector diffusion excitation vector

4

components by reading diffusion excitation vectors from the convolution section and multiplying said diffusion excitation vectors by decoder gains of said diffusion excitation
5 vectors, and a function which generates an excitation vector by adding said excitation vector adaptive code book components to said excitation vector diffusion excitation vector components and outputs it to said adaptive code
10 book; and an LPC synthesizing section having a function in which LPC-synthesizes said excitation vector.

[Claim 3] A speech decoder as set forth in Claim 1 or 2, wherein information to
15 generate the same code vectors as an algebraic code book is stored in an excitation storage section.

[Claim 4] A speech coder comprising a diffusion vector storage section for storing
20 either of random matrices, matrices obtained by learning or matrices obtained by knowledge; a digital filter; and a convolution section having functions which convolute the matrices stored in said diffusion vector storage section
25 to impulse responses or coefficients of said digital filter and output new impulse responses or new coefficients of said digital filter.

[Claim 5] A speech coder as set forth in Claim 4, further comprising an excitation vector information storage section for storing excitation vectors, wherein said convolution section convolutes excitation vectors produced on the basis of said excitation vectors and matrices stored in the diffusion vector storage section and also outputs diffusion excitation vectors.

10 [Claim 6] A speech coder as set forth in Claim 5, wherein information to generate the same code vectors as an algebraic code book is stored in an excitation vector storage section.

[DETAILED DESCRIPTION OF THE INVENTION]

15 [0001]

[Technical field of the invention] The present invention relates to an efficient compression decoding and coding of speech information in a speech decoder and a speech
20 coder.

[0002]

[Prior Art] In order to manage the increase in the number of subscribers in the field of movable communications such as digital
25 portable telephones, etc., a speech compression coding system of a low bit rate is demanded. In Japan, speech coding systems such as VSELP and

PSI-CELP are, respectively, employed and utilized as speech coding standardization system of digital portable telephones of full rate and half rate. Internationally, a coding system termed CS-ACELP is employed as an international standard speech coding system of 8kbps and is recommended by ITU-T recommendation G.729 (1995). These speech coding systems are those for which the CELP system (is described in "Code-Excited Linear Prediction (CELP): High Quality Speech At Very Low Rates" Manfred R. Schroeder. Bishnu S. Atal, Proc. ICASSP'85, pp.937-940) has been improved.

15 [0003]

Herein, a description is given of the basic algorithm of the CELP system. The CELP system is a system in which speech information is divided into excitation information and vocal track information and is coded. The system is featured in that such a system (A-b-S: Analysis by Synthesis) is employed, whereby the excitation information is coded by indices of a plurality of code vectors stored in a code book, and the vocal track information is coded by LPC (Linear Prediction Coefficients), wherein a comparison with input speech is

carried out by taking the vocal track information into consideration when coding the excitation information. Further, in the CELP, generally, the input speech is divided in terms
5 of a certain time period (a so-called "frame") and analyzed with respect to the LCP frame by frame, wherein the frame is divided into further smaller divisions (so-called a "sub frame"), and excitation vectors are searched
10 with respect to an adaptive code book and a probability code book.

[0004]

FIG. 5 is a functional block diagram of a prior art CELP type speech decoder. A parameter
15 decoder 152 obtains, through a transmitter 151, speech codes (LPC codes, probability code book indices, probability code book coding gains, adaptive code book indices, and adaptive code book coding gains) transmitted from a prior art
20 CELP type speech decoder (FIG. 6, which is described later). Next, it obtains a decoded LPC coefficient by decoding the LPC codes, obtains a decoding gain of the probability code book by decoding the coding gains of the
25 probability code book, and obtains a decoding gain of the adaptive code book by decoding the coding gain of the adaptive code book. Further,

the parameter decoder 152 outputs probability code book indices, probability code book decoding gains, adaptive code book indices, and adaptive code book decoding gains to the
5 excitation vector generator 155, and outputs the decoded LPC coefficients to the LPC synthesizer 156.

[0005]

The excitation vector generator 155 reads
10 adaptive code vectors based on the adaptive code book indices from the adaptive code book 153, and obtains excitation vector adaptive code book components by multiplying the obtained adaptive code vectors by the decoding
15 gains of the adaptive code book. Next, the generator 155 reads probability code vectors based on the probability code book indices from the probability code book 154 and obtains excitation vector probability code book
20 components by multiplying the obtained probability code vectors by the decoding gains of the probability code book. Further, the generator 155 adds the excitation vector adaptive code book components to the excitation
25 vector probability code book components to obtain excitation vectors, and outputs the obtained excitation vectors to the LPC

9

synthesizer 156 and the adaptive code book 153. Herein, previously stored code vectors in the adaptive code book 153 are renewed by the abovementioned excitation vectors inputted from
5 the excitation vector generator 155.

[0006]

The LPC synthesizer 156 carries out LPC synthesis on the excitation vectors obtained by the excitation vector generator 155 on the
10 basis of decoding LPC coefficients obtained by the parameter decoder 152, and the output thereof is transmitted to an output portion as digital output speech 157.

[0007]

15 FIG. 6 is a functional block diagram of a prior art CELP type speech coder. The LPC analyzing section 112 calculates an LPC coefficient by performing a self correlation analysis and a linear prediction analysis with
20 respect to frames existing in a digital input speech 111, and outputs LPC codes to a parameter coder 123 after it obtains the LPC codes by quantizing the LPC coefficients. And, the LPC analyzing section 112 obtains a
25 decoding LPC coefficient by decoding the LPC codes and next obtains impulse responses of a perceptual weighing filter having features such

as pitch weighting and high frequency weighting and outputs these to a perceptual weighting section 113. At the same time, the LPC analyzing section 112 further obtains impulse responses of the perceptual weighting LPC synthesis filter, and outputs these to a perceptual weighted LPC reverse synthesis filter A114, a perceptual weighted LPC synthesis section A116, a perceptual weighting LPC reverse synthesizer B119, and a perceptual weighting LPC synthesizer B121.

[0008]

A perceptual weighting section 113 filters inputted speech data subframe by subframe for weighting and subtracts a zero input response of the perceptual weighting LPC synthesis filter from the result of the output, whereby the perceptual weighting section 113 obtains target signals to be referred to when searching excitation vectors of an adaptive code book and outputs the target signals to the perceptual weighted LPC reverse synthesis filter A114 and subtracted 118.

[0009]

The perceptual weighted LPC reverse synthesis filter A114 reverses the target signals obtained by the perceptual weighting

section 113 in terms of time, synthesizes the obtained time-reversed signals by the perceptual weighting LPC synthesis filter having impulse responses given by the LPC analyzing section 112 as coefficients, reverses the output signals again in terms of time, and outputs them to a comparator A117 as the time-reverse synthesized outputs of the target signals.

10 [0010]

The adaptive code book 115 stores the previously stored excitation vectors received by the adaptive code book 115, and the previously stored excitation vector information is referred to as adaptive code vectors by the perceptual weighting LPC synthesizer A116, comparator A117, and adaptive code book renewing section 124.

[0011]

20 The perceptual weighting LPC synthesizer A116 reads adaptive code vectors from the adaptive code book 115, synthesizes the read adaptive code vectors by a perceptual weighting LPC synthesis filter having impulse responses obtained by the LPC analyzing section 112 as coefficients, and outputs the results by a comparator A117.

[0012]

First, the comparator A117 acquires the square value of the inner product of the adaptive code vector directly read from the adaptive code book 115 and the time-reverse synthesis output of the target signals obtained by the perceptual weighting LPC reverse synthesizer A114, and next acquires power of signals received from the perceptual weighting LPC synthesizer A116, in which a perceptual weighting LPC synthesis is applied to the adaptive code vectors, wherein the reference value of the adaptive code book retrieval is obtained by dividing the square value of the inner product by the power, indices of the adaptive code vectors read when the reference value is maximized, and the optimal gain to be multiplied by the code vectors are calculated, and these are outputted to the subtracter 118 and parameter coder 123. This series of processes is called an adaptive code book retrieval.

[0013]

The subtracter 118 subtracts a signal, which is obtained by multiplying an output signal, in which a code vector retrieved by an adaptive code book retrieval is synthesized in

terms of perceptual weighting LPC, by a gain from the target signal obtained by a perceptual weighting section 113, and the results of the subtraction are outputted to a perceptual
5 weighting LPC reverse synthesis section B119 as the target signal referred to when retrieving the probability code book.

[0014]

The perceptual weighting LPC reverse
10 synthesis section B119 reverses, in terms of time, the target signal for excitation retrieval of the probability code book generated in the subtracter 118, synthesizes it in terms of the perceptual weighting LPC,
15 reverses the output signal again in terms of time, and obtains the time reverse synthesized output of the target signal for excitation retrieval of the probability code book. Then, the synthesis section B119 outputs the output
20 thereof to the comparator B122.

[0015]

The probability code book 120 stores a plurality of code vectors, and these code vectors are referred to as probability code
25 vectors by the perceptual weighting LPC synthesis section B121, comparator B122, and adaptive code book renewing section 124.

[0016]

The perceptual weighting LPC synthesis section B121 synthesizes the probability code vectors read from the probability code book 120 by a perceptual weighting LPC synthesizing filter having impulse responses obtained from the LPC analyzing section 112 as coefficients, and outputs the synthesized signal to the comparator B122.

10 [0017]

As shown in Expression (1), the comparator B122 first obtains a square value of the inner product of the i th probability code vector $V(i,n)$ directly read from the probability code book 120 and a time reverse synthesized output $r(n)$ of the target signal obtained by the perceptual weighting LPC reverse synthesis section B119, and, next, calculates the power of the synthesized signal $S(i,n)$ received by the perceptual weighting LPC synthesis section B121. By dividing the square value of the inner product by the power, the standard value $std(i)$ of the probability code book is obtained, wherein an index to express the number of probability code vectors read when the standard value is maximized, the optimal gain to be multiplied by the

probability code vectors are calculated, and the results are outputted to a parameter coder 123. These series of processes are called a "probability code book".

5 [0018]

[Expression (1)]

$$std(i) = \frac{\left| \sum_{n=0}^{N-1} V(i,n) \cdot r(n) \right|^2}{\sum_{n=0}^{N-1} S(i,n)}$$

10

where std(i): standard value for probability code book retrieval, i: index of the code vector, N: length of a subframe, n: component number ($0 \leq n(N)$) of a vector, V(i,n): ith probability code vector, S(i,n): vector in which the V(i,n) is synthesized in terms of perceptual weighting LPC, and r(n): time reverse synthesized output vector of the target signal for probability code book retrieval.

20 [0019]

First, the parameter coder 123 codes the optimal gain to be multiplied by an adaptive code vector obtained from the comparator A117 and the optimal gain to be multiplied by the probability code vector obtained by the comparator B122 in order to obtain the coded gain of the adaptive code book and that of the

probability code book. Next, the obtained coded gain of the adaptive code book and that of the probability code book are decoded, whereby the decoded gain of the adaptive code book and that of the probability code book are, respectively, obtained, and further the code gain of the adaptive code book, code gain of the probability code book, index of the adaptive code book, index of the probability code book and LPC code are outputted to the transmitter 125. Then, the decoded gain of the adaptive code book, decoded gain of the probability code book, index of the adaptive code book, and index of the probability code book are outputted to the adaptive code book renewing section 124.

[0020]

The adaptive code book renewing section 124 receives input from the parameter coder 123, reads an adaptive code vector based on the index of the adaptive code book from the adaptive code book 115, obtains excitation vector adaptive code book components by multiplying the read adaptive code vector by the decoded gain of the adaptive code book, next reads a probability code vector based on the index of the probability code book from the

probability code book 120, and obtains
 excitation vector probability code book
 components by multiplying the rear probability
 code vector by the decoded gain of the
 5 probability code book, wherein the excitation
 vector adaptive code book components are added
 to the excitation vector probability code book
 components in order to obtain excitation
 vectors, and the excitation vectors are
 10 outputted to the adaptive code book 115.
 Herein, the previously stored code vectors
 existing in the adaptive code book 115 are
 renewed by the abovementioned excitation
 vectors inputted by the adaptive code book
 15 renewing section 124.

[0021]

As the abovementioned probability code
 book used for the CELP, there are a random code
 book and an algebraic code book (written in
 20 "8KBIT/S ACELP CODING OF SPEECH WITH 10MS
 SPEECH-FRAME: A CANDIDATE FORCCITT
 STANDARDIZATION": R.Salami, C.Laflamme, J-P.
 Adoul, ICASSP' 94, pp.II-97-II-100, 1994), etc.
 A brief description is given of each of the
 25 above cases.

[0022]

"Random code book" is one of the most

classical code books and stores random matrices prepared by the random numbers. Since the code book is random in character, it is possible to obtain high quality synthesized sound. But, since all code vectors are stored in advance, a large ROM capacity is required. Further, since the code book retrieval accompanies perceptual weighting LPC synthesis with all code vectors, a great number of calculations are required.

10 [0023]

The individual code vectors of the "algebraic code book" are featured in that it consists of four pulses, the size of which is 1 (amplitude is +1 or -1) and the position of the respective pulses is determined by calculations of the indices. Therefore, no ROM is required for a code book. But, since a small number of pulses are sufficient, the sound quality of the synthesized sound is worsened (especially, an unvoiced section). And, the greatest feature of the algebraic code book resides in that a code book retrieval is enabled with only a small number of calculations. This feature of the algebraic code book resides in that, since only four pulses, the size of which is 1, are provided in $V(i,n)$, a numerator of the expression (1) can

be calculated by adding the values of four samples in $r(n)$ (addition: three times) and squaring the result of the addition (multiplication: one time), and if a self-correlation matrix of impulses of the perceptual weighting LPC synthesis filter, which are obtained in advance, is stored in a RAM, the denominator in the expression (1) can be calculated by additions of 15 times at most (the amount of actual calculation can be further decreased if a quadruple loop character and symmetry of self-correlation matrix are utilized).

[0024]

15 [Problems to be solved by the invention]

If the random code book is used, since a perceptual weighting LPC synthesis filter can be driven with excitation having random characters, the sound quality of the synthesized speech is improved. However, the amount of calculations of the numerator and denominator in the expression (1) is increased, whereby the amount of the code book retrieval is also increased, and the capacity of a ROM in which all code vectors are stored in advance will be increased.

[0025]

On the other hand, if the algebraic code book is used, the amount of calculations of the numerator and denominator in the expression (1) can be reduced, and it is not necessary for the code vectors to be stored as they are, wherein the ROM capacity may be reduced. However, since the perceptual weighting LPC synthesis filter is driven by a small number of pulses, the sound quality of the synthesized sound (especially an unvoiced section) is worsened.

[0026]

It is therefore an object of the present invention to provide a speech decoder and speech coder, having advantageous features, in which the amount of calculations is reduced for a code book retrieval, the ROM capacity is reduced, and high quality synthesized sound can be provided.

[0027]

[Means for Solving the Problems]

In order to solve the themes, a speech decoder according to the invention comprises an excitation information storing section for storing excitation information; a diffusion vector storing section for storing either matrix of random matrices, matrices obtained by learning, and matrices obtained by knowledge;

and a convolution section, for outputting diffusion excitation vectors, which convolutes excitation vectors generated on the basis of said excitation information and matrices stored
5 in said diffusion vector storing section, wherein excitation vector generation information constituted by only a small number of pulses is stored in an excitation information storage section, either of the
10 matrices described above is stored in a diffusion vector storage section, and at the same time, diffusion excitation vectors are constituted by convoluting excitation information in the excitation information
15 storage section and diffusion vectors in the diffusion vector storage vector.

[0028]

Thereby, it is possible to obtain a speech decoder having excellent synthesized
20 speech quality (especially, the unvoiced section) and not requiring any random code book, for which only small ROM capacity is sufficient.

[0029]

25 Further, a speech coder according to the invention comprises an adaptive code book for storing the previously stored excitation

vectors; an excitation vector generator having a function to generate excitation vector adaptive code book components by reading adaptive code vectors from said adaptive code book and multiplying said adaptive code vectors by decoder gains of said adaptive code book, a function to generate excitation vector diffusion excitation vector components by reading diffusion excitation vectors from the convolution section and multiplying said diffusion excitation vectors by decoder gains of said diffusion excitation vectors, and a function which generates an excitation vector by adding said excitation vector adaptive code book components to said excitation vector diffusion excitation vector components and outputs it to said adaptive code book; and an LPC synthesizing section having a function in which LPC-synthesizes said excitation vector, wherein a new pulse response to determine new filter characteristics and a new coefficient are prepared by convoluting an impulse response or coefficient in the digital filter and a diffusion vector in the diffusion vector storage section.

[0030]

Thereby, it is possible to obtain a

speech coder which enables code book retrieval by only a small number of calculations as in the retrieval of excitation vectors, which is constituted by a small number of pulses.

5 [0031]

[Embodiments of the invention]

The first aspect of the invention as set forth in Claim 1 of the invention is a speech decoder featured in comprising an excitation
10 information storing section for storing excitation information; a diffusion vector storing section for storing either random matrices, matrices obtained by learning, and matrices obtained by knowledge; and a
15 convolution section, for outputting diffusion excitation vectors, which convolutes excitation vectors generated on the basis of said excitation information and matrices stored in said diffusion vector storing section. These
20 pertain to solving the themes related to transmission of speech information. The speech decoder need not have any probability code book such as a random code book and/or an algebraic code book, whereby the ROM capacity can be
25 reduced, the number of pulses of speech information can be reduced, and the amount of calculations can be reduced.

[0032]

The second aspect of the invention as set forth in Claim 2 is a speech decoder as set forth in Claim 1, wherein the speech decoder is
5 featured in comprising an adaptive code book for storing the previously stored excitation vectors; an excitation vector generator having a function to generate excitation vector adaptive code book components by reading
10 adaptive code vectors from said adaptive code book and multiplying said adaptive code vectors by decoder gains of said adaptive code book, a function to generate excitation vector diffusion excitation vector components by
15 reading diffusion excitation vectors from the convolution section and multiplying said diffusion excitation vectors by decoder gains of said diffusion excitation vectors, and a function which generates an excitation vector
20 by adding said excitation vector adaptive code book components to said excitation vector diffusion excitation vector components and outputs it to said adaptive code book; and an LPC synthesizing section having a function in
25 which LPC-synthesizes said excitation vector, wherein vocal track information of speech information can be synthesized in terms of LPC

to generate good quality synthesized sound, a probability code book such as a random code book and/or an algebraic code book, etc., is not required, the ROM capacity can be reduced, 5 the number of pulses of excitation information can be reduced, and the amount of calculations can also be further reduced.

[0033]

The third aspect of the invention as set forth in Claim 3 is a speech decoder as set 10 forth in Claim 1 or 2, which is featured in that information to generate the same code vectors as those of the algebraic code book is stored in the excitation information storage 15 section, and has further actions in that the number of pulses of excitation information can be reduced and the amount of calculations can also be reduced.

[0034]

20 The fourth aspect of the invention as set forth in Claim 4 is a speech coder featured in comprising a diffusion vector storage section for storing either of random matrices, matrices obtained by learning or matrices obtained by 25 knowledge; a digital filter; and a convolution section having functions which convolute the matrices stored in said diffusion vector

storage section to impulse responses or coefficients of said digital filter and output new impulse responses or new coefficients of said digital filter. That is, the speech coder
5 relates to solving the themes pertaining to transmission of speech information, wherein new impulse responses to determine new filter characteristics and new coefficients are prepared by convoluting the impulse responses
10 or coefficients in a digital filter and diffusion vectors in a diffusion vector storage section, and a digital filter is formed, which is necessary to code in compliance with decoding in the speech decoder.

15 [0035]

The fifth aspect of the invention as set forth in Claim 5 is a speech coder featured in further comprising an excitation vector information storage section for storing
20 excitation vectors, wherein said convolution section convolutes excitation vectors produced on the basis of said excitation vectors and matrices stored in the diffusion vector storage section and also outputs diffusion excitation
25 vectors, wherein new impulse responses to determine new filter characteristics and new coefficients are prepared by convoluting the

impulse responses or coefficients in a digital filter and diffusion vectors in a diffusion vector storage section, and a digital filter is formed, which is necessary to code in compliance with decoding in the speech decoder, and wherein it is not necessary to provide any probability code book such as a random code book, an algebraic code book, etc., the ROM capacity can be reduced, the number of pulses in speech information can also be reduced, and minimal calculations is sufficient.

[0036]

The sixth aspect of the invention as set forth in Claim 6 is a speech coder as set forth in Claim 5, wherein information to generate the same code vectors as an algebraic code book is stored in an excitation vector storage section, and wherein the number of pulses in excitation information can be reduced, and only minimal calculations are sufficient.

[0037]

Hereinafter, embodiments of the present invention are explained with reference to FIG. 1 and FIG. 4.

(First embodiment) One embodiment of the present invention is explained with reference to FIG. 1 and FIG. 2.

[0038]

FIG. 1 is a block diagram showing the major parts of a speech decoder according to the invention. In FIG. 1, an excitation information storage section is indicated by 1, a diffusion vector storage section by 2, and a convolution section by 3. The excitation information storage section 1 stores excitation information of generation information, etc., of the same code vectors as those in, for example, an algebraic code book, and outputs the excitation information as excitation vectors. The diffusion vector storage section 2 stores, for example, diffusion information such as random matrices, etc., and outputs the diffusion information as diffusion vectors. The convolution section 3 inputs and convolutes excitation vectors and diffusion vectors, and outputs as diffusion excitation vectors.

[0039]

FIG. 2 shows a CELP type speech decoder in which the major parts of the speech decoder shown in FIG. 1 are used instead of a prior art probability code book. The speech information storage section 1 stores generation information of the same code vectors as those in the algebraic code book, and the diffusion vector

storage section 2 stores random matrices as diffusion vectors. A parameter decoder 4 first acquires speech codes (LPC codes, diffusion speech index, coded gain of the diffusion excitation vectors, indices of the adaptive code book and coded gain of the adaptive code book) through a transmitter 5 and obtains a decoded LPC coefficient by decoding the LPC codes. Next, a decoded gain of the diffusion excitation vector is obtained by decoding the coded gain of the diffusion excitation vector, and a decoded gain of the adaptive code book is obtained by decoding the coded gain of the adaptive code book. Further, the diffusion excitation index is outputted to the convolution section 3, and the index and decoded gain of the adaptive code book and the decoded gain of the diffusion excitation vector are outputted to an excitation vector generator 7, wherein the decoded LPC coefficient is outputted to an LPC synthesizer 8.

[0040]

The convolution section 3 reads excitation vectors based on the diffusion excitation index from the excitation information storage section 1, next reads the diffusion vectors stored in the diffusion

vector storage section 2, further generates diffusion excitation vectors by convoluting the read excitation vectors and diffusion vectors, and finally outputs to the excitation vector generator 7. The excitation vector generator 7 first reads adaptive code vectors based on the index of the adaptive code book, acquires excitation vector adaptive code book components by multiplying the obtained adaptive code vector by a decoded gain of the adaptive code book, and further obtains excitation vector diffusion excitation vector components by multiplying the diffusion excitation vector obtained from the convolution section 3 by the decoded gain of the diffusion excitation vector. Next, excitation vectors are acquired by adding the obtained excitation vector adaptive code book components and the excitation vector diffusion excitation vector components to each other, wherein the obtained excitation vectors are outputted to the LPC synthesizer 8 and adaptive code book 6. Herein, the previously stored code vectors existing in the adaptive code book 6 are renewed by the abovementioned excitation vectors inputted from the excitation vector generator 7. The LPC synthesizer 8 synthesizes the excitation

vectors inputted from the excitation vector generator 7 by an LPC synthesis filter having a decoded LPC coefficient acquired from the parameter decoder 4, and obtains output speech.

5 [0041]

Further, since, in the preferred embodiment, the diffusion vector storage section 2 stores random matrices as diffusion vectors, the abovementioned excitation vector
10 diffusion excitation vector components will have random characteristics. The sound quality of the output speech (especially, unvoiced section) can be improved. Further, the diffusion excitation vectors can be generated
15 by excitation information of the excitation information storage section 1 and diffusion vectors of the diffusion vector storage section 2. Therefore, the ROM capacity can be decreased.

20 [0042]

Also, although, in the preferred embodiment, a description was given of an example in which generation information of the same code vectors as in the algebraic code book
25 is stored in the excitation information storage section 1, the preferred embodiment may be applicable to generation information of the

other code books or in a case where the other code book itself is stored. Also, in the preferred embodiment, a description was given of an example in which random matrices are stored in the diffusion vector storage section 2. However, the embodiment may be applicable to cases where matrices obtained by other learning or matrices obtained by knowledge are used.

[0043]

10 A CELP type speech decoder is used for the preferred embodiment. However, the invention may be applicable to VOCODER type, or other speech decoders.

[0044]

15 (Second embodiment) One embodiment of the present invention is explained with reference to FIG. 3 and FIG. 4.

[0045]

FIG. 3 is a block diagram of the major parts of a speech coder according to the invention. In FIG. 3, a diffusion vector storage section is indicated by 11, a digital filter is indicated by 12, a convolution section is indicated by 13, and a speech information storage section is indicated by 14. The diffusion vector storage section 11 stores diffusion information such as, for example,

random matrices, and outputs the diffusion information as diffusion outputs. The digital filter 12 filters and outputs input signals, and at the same time, outputs filter information which determines the characteristics of the filter itself. In FIG. 3, output of impulse responses or coefficients is shown as filter information. The excitation information storage section 14 stores excitation information of generation information, etc., of the same code vectors as in, for example, an algebraic code book, and outputs the excitation information as excitation vectors. In FIG. 3(a), the convolution section 13 inputs and convolutes diffusion vectors and impulse responses or coefficients and outputs new impulse responses or new coefficients. In FIG. 3(b), in addition to the features shown in FIG. 3(a), the convolution section 13 inputs and convolutes excitation vectors and diffusion vectors, and outputs as diffusion excitation vectors.

[0046]

FIG. 4 shows a CELP type speech coder in which the main parts of the speech coder shown in FIG. 3 are used instead of a prior art probability code book. The excitation

information storage section 14 stores generation information of the same code vectors as in the algebraic code book, and the diffusion vector storage section 11 stores
5 random matrices as diffusion vectors. Input speech is indicated by 15, which is digital input speech data. The LPC analyzing section 16 calculates an LPC coefficient by carrying out a self-correlation analysis and a linear
10 prediction analysis with respect to a certain frame in the input speech 5, quantizes the LPC coefficient in order to obtain an LPC code, and outputs it to the parameter coder 17, whereby the LPC code is decoded to obtain a decoded LPC
15 coefficient. Next, an impulse response of the perceptual weighting filter having such characteristics as pitch enhancement and high frequency enhancement, etc., is obtained and outputted to the perceptual weighting section
20 18, and at the same time an impulse response of the perceptual weighting LPC synthesis filter is obtained and is outputted to the perceptual weighting LPC reverse synthesis section A19 and perceptual weighting LPC synthesis section A20.
25 Either of the perceptual weighting LPC reverse synthesis section A19 or the perceptual weighting LPC synthesis section A20 contains a

digital filter.

[0047]

The perceptual weighting section 18 carries out a perceptual weighting filtering subframe by subframe with respect to the input speech data, and the perceptual weighting section 18 subtracts a zero input response of the perceptual weighting LPC synthesis filter from the results thereof in order to obtain the target signal to be referred to when retrieving an excitation of the adaptive code book, and outputs the target signal to the perceptual weighting LPC reverse synthesis section A19 and the subtracter 18. The perceptual weighting LPC reverse synthesis section A19 reverses the target signal inputted from the perceptual weighting section 18 in terms of time, and the reversed signal is synthesized by a perceptual weighting LPC synthesis filter having an impulse response given from the LPC analyzing section 16 as a coefficient, wherein the outputted signal is reversed again in terms of time in order to obtain time-reversed synthesis output of the target signal, and the output is given to the comparator A21.

[0048]

The adaptive code book 22 stores the

previously stored excitation vectors received from the adaptive code book renewing section 23, and the previously stored excitation vector information is referred to as adaptive code
5 vectors by the perceptual weighting LPC synthesis section A20, comparator A21, and adaptive code book renewing section 23. The perceptual weighting LPC synthesis section A20 first reads the adaptive code vectors from the
10 adaptive code book 22, synthesizes the read adaptive code vectors by a perceptual weighting LPC synthesis filter having an impulse response obtained from the LPC analyzing section 16 as a coefficient, and outputs the results to the
15 comparator A21. Next, the impulse response of the perceptual weighting LPC synthesis filter is outputted to the convolution section 13.

[0049]

First, the comparator A21 obtains the
20 square value of the inner product of the adaptive code vectors directly read from the adaptive code book 22 and the time-reversed synthesis output of the target signal obtained by the perceptual weighting LPC reverse
25 synthesis section A19. Next, the comparator A21 obtains power of a signal in which a perceptual weighting LPC synthesis is given to the

adaptive code vectors received from the perceptual weighting LPC synthesis section A20, wherein by dividing the abovementioned square value of the inner product by this power, the standard value of an adaptive code vector retrieval is obtained. The index of the adaptive code vector read when the standard value is maximized, and the optimal gain to be multiplied by the adaptive code vectors are calculated, and outputted to the subtracter 24 and parameter coder 7.

[0050]

The subtracter 24 synthesizes the adaptive code vector retrieved in the adaptive code book retrieval from the target signal obtained in the perceptual weighting section 18 in terms of perceptual weighting LPC, and subtracts a signal which is obtained by multiplying the synthesized signal by the optimal gain of the adaptive code book obtained in the comparator A21. The subtracted result is outputted to the perceptual weighting LPC reverse synthesis section B25 as the target signal to be referred to when retrieving a diffusion excitation vector.

[0051]

The diffusion vector storage section 11

stores random matrices. The convolution section 13 performs a convolution operation of the impulse response of the perceptual weighting LPC synthesis filter, which is received from the perceptual weighting LPC synthesis section A20, and the diffusion vector read from the diffusion vector storage section 11. The results of the convolution operation are outputted to the perceptual LPC reverse synthesis section B25 and the perceptual weighting LPC synthesis section B26 as a new impulse response. Either of the perceptual weighting LPC reverse synthesis section B25 or the perceptual weighting LPC synthesis section B26 includes a digital filter.

[0052]

The perceptual weighting LPC reverse synthesis section B25 reverses, in terms of time, the target signal obtained when retrieving a diffusion excitation vector generated by the subtracter 24, and the reversed signal is synthesized by a perceptual weighting LPC synthesis filter having a new impulse response obtained from the convolution section 13 as a coefficient. The output signal is reversed again in terms of time, and outputted to the comparator section B27 as a

time-reversed synthesis output of the target signal.

[0053]

Since the excitation information storage
5 section 14 stores generation information of the
same code vectors as in the algebraic code
book, code vectors consisting of a small number
of pulses (herein, four pulses) can be
generated. The generated code vectors are
10 referred to as excitation vectors by the
perceptual weighting LPC synthesis section B26,
comparator B27 and convolution section 13. The
perceptual weighting LPC synthesis section B26
synthesizes the excitation vectors read from
15 the excitation information storage section 14
by the perceptual weighting LPC synthesis
filter having a new impulse response read from
the convolution section 13 as a coefficient.
And, a self-correlation matrix of the
20 synthesized signal is obtained and outputted to
the comparator B27.

[0054]

First, the comparator B27 obtains the
square value of the inner product of the
25 excitation vector directly read from the
excitation information storage section 14 and
the time-reversed synthesis output of the

target signal obtained by the perceptual weighting LPC reverse synthesis section B25. Next, with reference to the self-correlation matrices reversed by the perceptual weighting
5 LPC synthesis section B26, power of the signal, in which the excitation vector is provided with perceptual weighting LPC synthesis, is obtained. And, by dividing the square value of the inner product by the power, the standard
10 value of the diffusion excitation vector retrieval is then obtained. A diffusion excitation vector expressing the number of the excitation vector read when the standard value is maximized, and the optimal gain to be
15 multiplied by the diffusion excitation vector are calculated, and outputted to the parameter coder 17.

[0055]

First, the parameter coder 17 codes the
20 optimal gain to be multiplied by the adaptive code vector obtained by the comparator A21 and the optimal gain to be multiplied by the diffusion excitation vector obtained by the comparator B27 in order to obtain a coded gain
25 of the adaptive code book and a coded gain of the diffusion excitation vector. Next, the obtained coded gain of the adaptive code book

and coded gain of the diffusion excitation vector are decoded to obtain the decoded gain of the adaptive code book and that of the diffusion excitation vector. Further, the code gain of the adaptive code book, coded gain of the diffusion excitation vector, LPC code, index of the adaptive code book, and index of the diffusion excitation are outputted to the transmitter 28, whereby the decoded gain of the adaptive code book and index of the adaptive code book are outputted to the adaptive code book renewing section 23, and the decoded gain of the diffusion excitation vector and the diffusion excitation index are outputted to the convolution section 13.

[0056]

First, the convolution section 13 reads an excitation vector based on the diffusion excitation index from the excitation information storage section 14, and next reads a diffusion vector stored in the diffusion vector storage section 11. Next, the read excitation vector and diffusion vector are convoluted to generate the diffusion excitation vector. An excitation vector diffusion excitation vector component is obtained by multiplying the diffusion excitation vector by

the decoded gain, and the results are outputted to the adaptive code book renewing section 23. The adaptive code book renewing section 23 first obtains an adaptive code vector based on the index of the adaptive code book received from the parameter coder 17 from the adaptive code book 22, and obtains excitation vector adaptive code book components by multiplying the obtained adaptive code vector by the decoded gain of the adaptive code book. Next, the obtained excitation adaptive code book components are added to the excitation vector diffusion excitation vector components inputted from the convolution section 13 in order to generate excitation vectors, wherein the generated excitation vectors are outputted to the adaptive code book 22. Herein, the previously stored code vectors existing in the adaptive code book 22 are renewed by the abovementioned excitation vectors inputted from the adaptive code book renewing section 23.

[0057]

Further, although, in the preferred embodiment, a description was given of an example in which generation information of the same code vectors as in the algebraic code book is stored in the excitation information storage

section 14, the preferred embodiment can be carried out as well in cases where generation information of the other code books or the other code book itself are stored. Furthermore, in the preferred embodiment, a description was given of an example where random matrices are stored in the diffusion vector storage section 11. However, the preferred embodiment may be applicable to cases where matrices obtained from learning or those obtained by knowledge are used.

[0058]

A CELP type speech decoder is used for the preferred embodiment. However, the invention may be applicable to VOCODER type, or other speech decoders.

[0059]

[Effects of the invention]

As described above, with the present invention, various advantages can be brought about, for example, minimal calculations may be sufficient for retrieval of a code book, the ROM capacity may be reduced, and excellent quality synthesized speech can be developed.

[BRIEF DESCRIPTION OF THE DRAWINGS]

[FIG. 1] is a block diagram showing the major parts of a speech decoder according to a

preferred embodiment of the invention,

[FIG. 2] is a block diagram showing a CELP type speech decoder according to the preferred embodiment of the invention,

5 [FIG.3](a) is a block diagram showing the major parts of a speech coder according to a preferred embodiment of the invention, and (b) is a block diagram showing the major parts of a speech coder according to a preferred
10 embodiment of the invention,

[FIG. 4] is a block diagram showing a CELP type speech coder according to a preferred embodiment of the invention,

[FIG. 5] is a block diagram showing a prior art
15 CELP type speech decoder, and

[FIG. 6] is a block diagram showing a prior art CELP type speech coder.

[Description of the Symbols]

- | | |
|------|--|
| 1 | Excitation information storage section |
| 20 2 | Diffusion vector storage section |
| 3 | Convolution section |
| 4 | Parameter decoder |
| 5 | Transmitter |
| 6 | Adaptive code book |
| 25 7 | drive excitation vector generator |
| 8 | LPC synthesis section |
| 9 | Output speech |

- 11 Diffusion vector storage
- 12 Digital filter
- 13 Convolution section
- 14 Excitation information storage section
- 5 15 Input speech
- 16 LPC analysis section
- 17 Parameter coder
- 18 Perceptual weighting section
- 19 Perceptual weighting LPC reverse synthesis
- 10 section A
- 20 Perceptual weighting LPC synthesis section
- A
- 21 Comparator A
- 22 Adaptive code book
- 15 23 Adaptive code book renewing section
- 24 Subtractor
- 25 Perceptual weighting LPC reverse synthesis
- section B
- 26 Perceptual weighting LPC synthesis section
- 20 B
- 27 Comparator B
- 28 Transmitter

FIG. 1

- 1 Excitation information storage
- Excitation vectors
- 3 Convolution section
- 5 Diffusion vectors
- 2 Diffusion vector storage section
- Diffusion excitation vector

FIG. 2

- 5 Transmitter
- 10 4 Parameter decoder
- 9 Output speech
- Adaptive code book
- Excitation information storage section
- 3 Convolution section
- 15 2 Diffusion vector storage section
- Excitation vector generator
- 1 LPC synthesis section

FIG. 5

- 151 Transmitter
- 20 153 Adaptive code book
- 154 Probability code book
- 152 Parameter decoder
- 155 Excitation vector generator
- 157 Output speech
- 25 156 LPC synthesis section

FIG. 3(a)

- 11 Diffusion vector storage section

Diffusion vectors
 12 Digital filter
 Impulse response/coefficient
 13 Convolution section
 5 New impulse response/New coefficient

FIG. 3(b)

12 Digital filter
 Impulse response/coefficient
 11 Diffusion vector storage section
 10 Diffusion vector
 Impulse response/coefficient
 13 Convolution section
 New impulse response/new coefficient
 Diffusion excitation vector
 15 Excitation vector
 14 Excitation information storage section

FIG. 4

Input speech
 LPC analysis section
 20 Perceptual weighting
 Transmitter
 Perceptual weighting LPC reverse synthesis
 section A
 Adaptive code book
 25 Perceptual weighting LPC synthesis section A
 Comparator A
 Parameter coder

Subtractor
Adaptive coder renewing section
Diffusion vector storage section
Convolution section
5 Perceptual weighting LPC reverse synthesis
section B
Excitation information storage section
Perceptual weighting LPC synthesis section B
Comparator B
10 FIG. 6
Input speech
Transmitter
LPC analysis section
Perceptual weighting
15 Perceptual weighting LPC reverse synthesis
section A
Adaptive coder
Perceptual weighting LPC synthesis section A
Comparator A
20 Parameter coder
Subtractor
Adaptive code book renewing section
Perceptual weighting LPC reverse synthesis
section B
25 Probability code book
Perceptual weighting LPC synthesis section B
Comparator B